

目標：インフラ構造物の定期点検をAI・ロボット技術で自動化！

背景

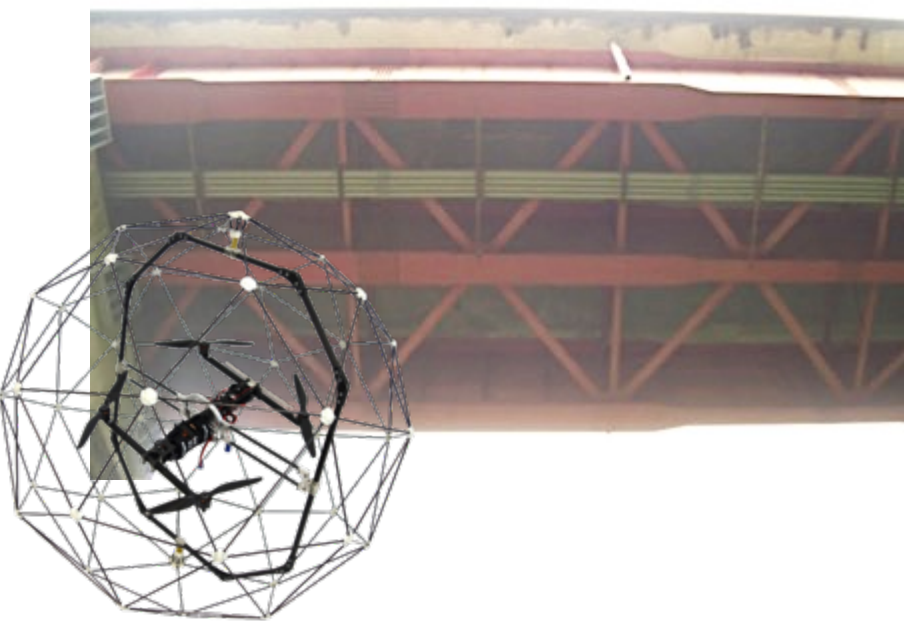
- 全国のインフラの老朽化が急速に進んでいる
- 点検員と予算の確保が年々難しくなっている

現在の点検



- 専用車両を使用
- 熟練者による点検
- 交通をストップ

目指す点検

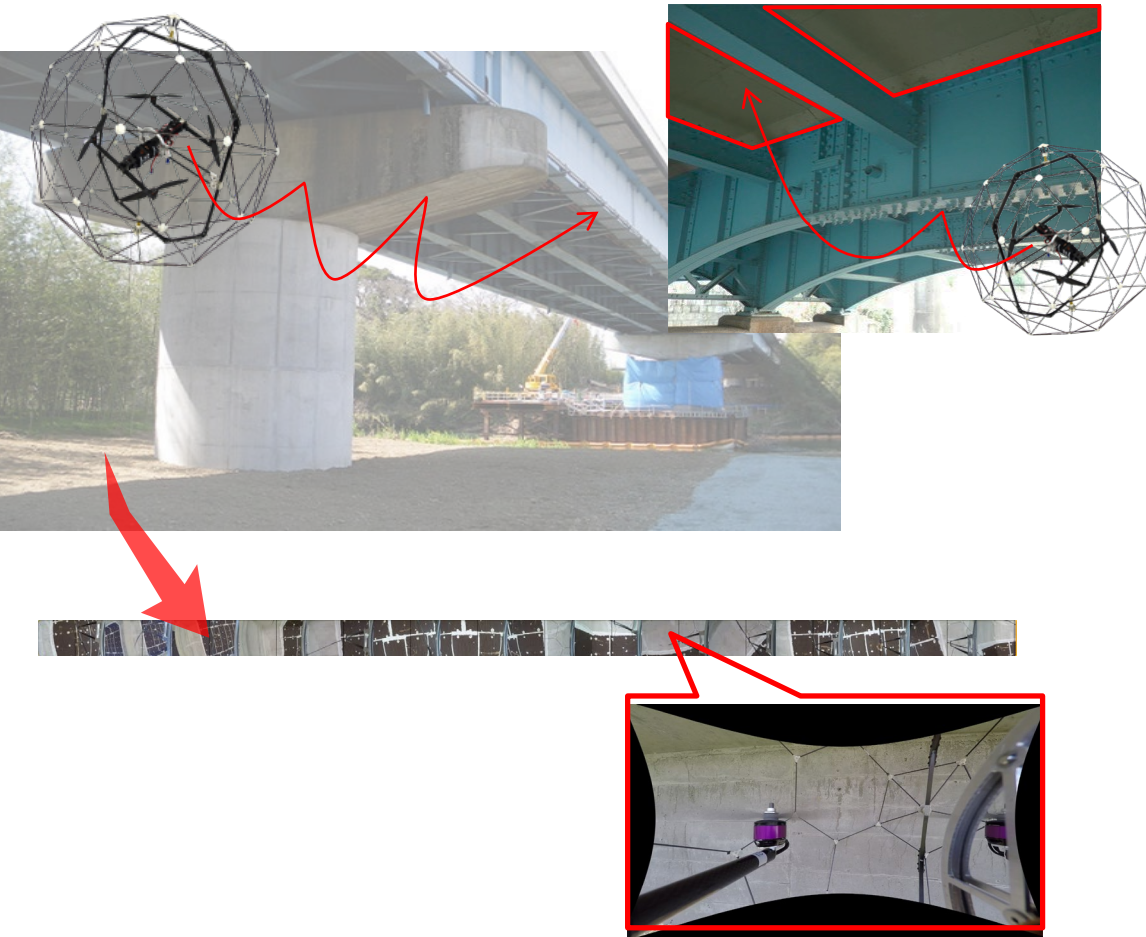


- ドローンが自律飛行
- AIが自動点検
- 交通を妨げない

点検開始から完了までのステップと課題

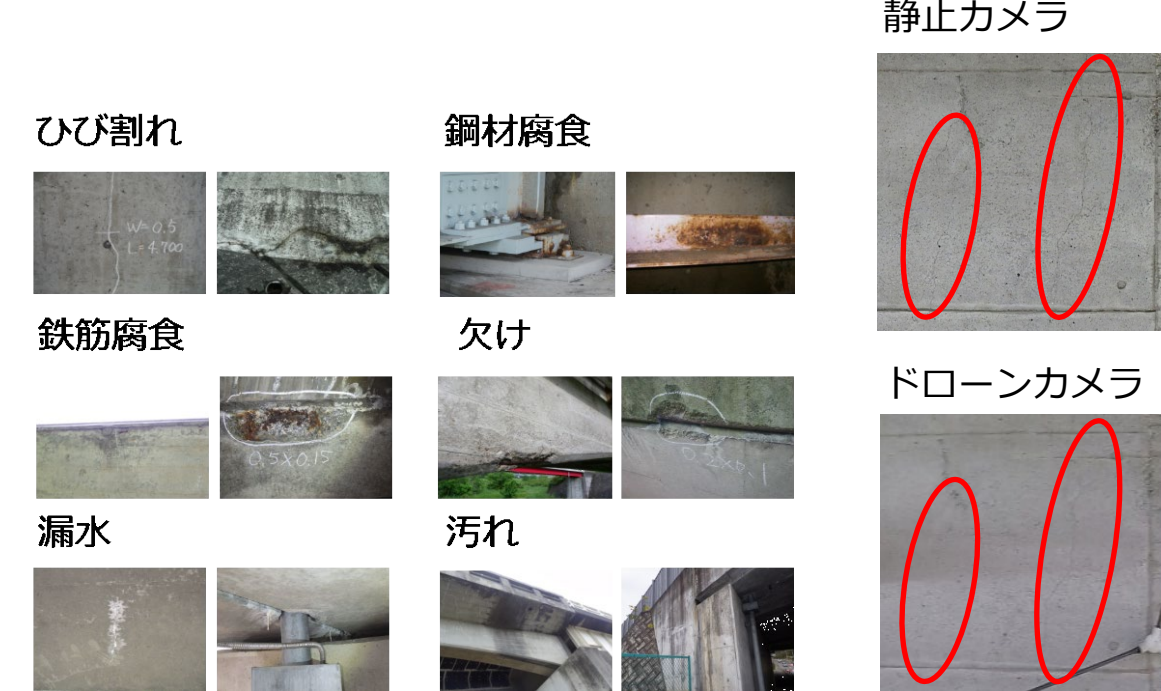
飛行

- 非GPS環境下での自律飛行
- 限定環境以外での成功例はまだない
- 事前の地図作成が不可能に近い
- 演算処理能力の制約（組み込み用CPU・GPUの限界）



画像による検査

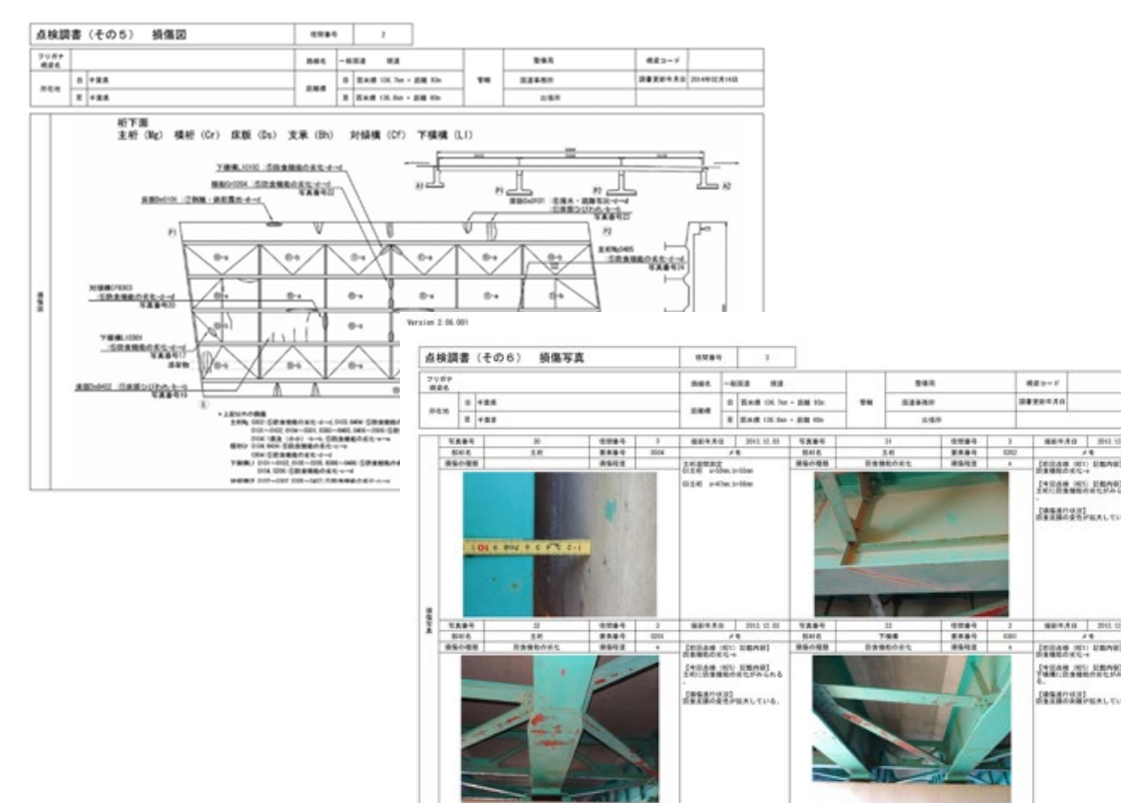
- 低画質画像からの損傷認識
- ドローンによる撮影⇒様々な画質の低下（動きによるブレ・照明不足によるセンサノイズ等）
- 少数サンプルからの学習
- 損傷は見た目の変動が大きいわりに、学習データは入手困難



レポート作成

(点検調書)

- 自然言語(文章)による説明文の生成
- 画像理解という難問へ挑戦

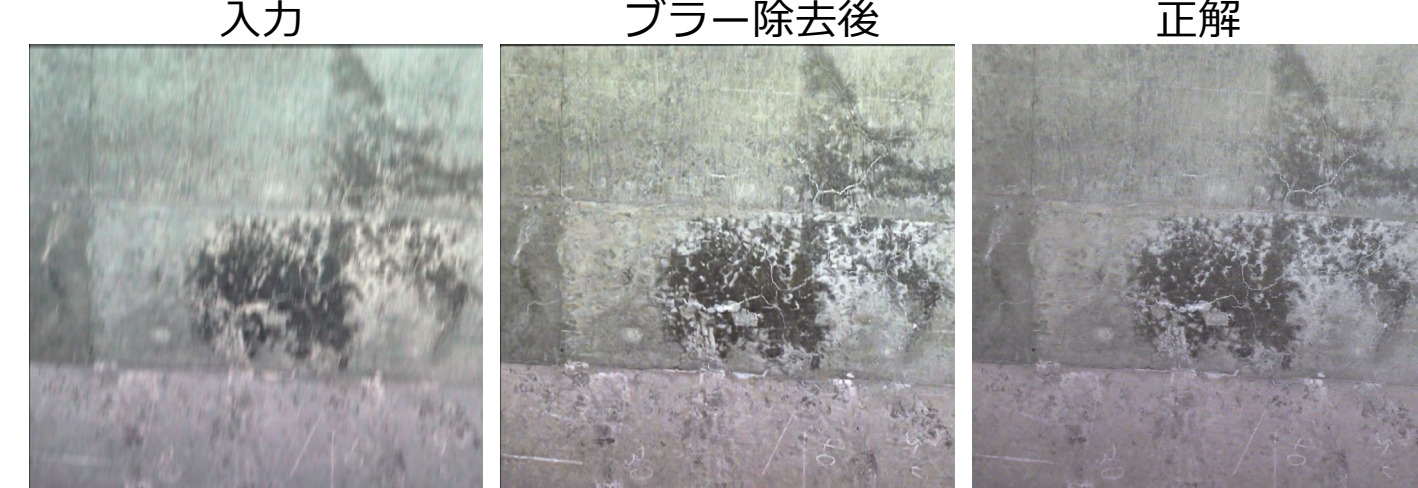


(注：見本です)

画質劣化の克服 [Liu-Sun-Suganuma-Okatani, CVPR2019]

- 小型ドローンによる画像撮影：画質の制約
- 撮像系の重さと画質にトレードオフ（機体サイズ・飛行時間に影響）
- 様々な画質劣化要因に有効な「ペア演算」を有効活用できる手法を開発
- 5つのタスクで世界最高性能を達成 ※ 2019年2月時点で

モーションブラー

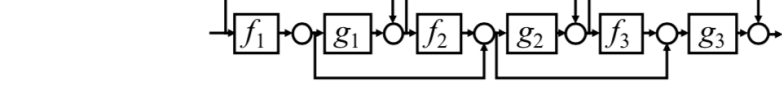


ノイズ

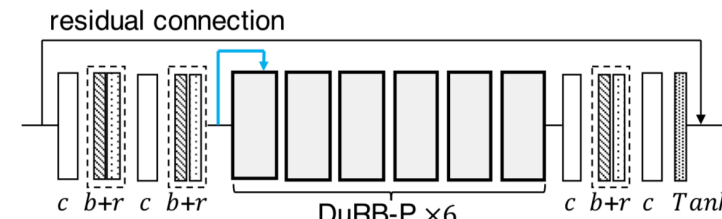


開発したネットワーク構造

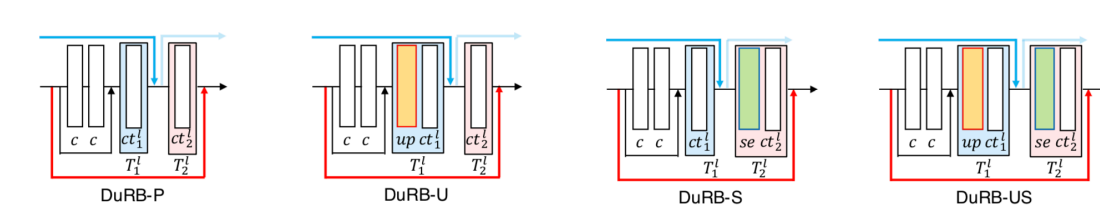
“Dual Residual Network”



全体構造



四種類の基本ブロック



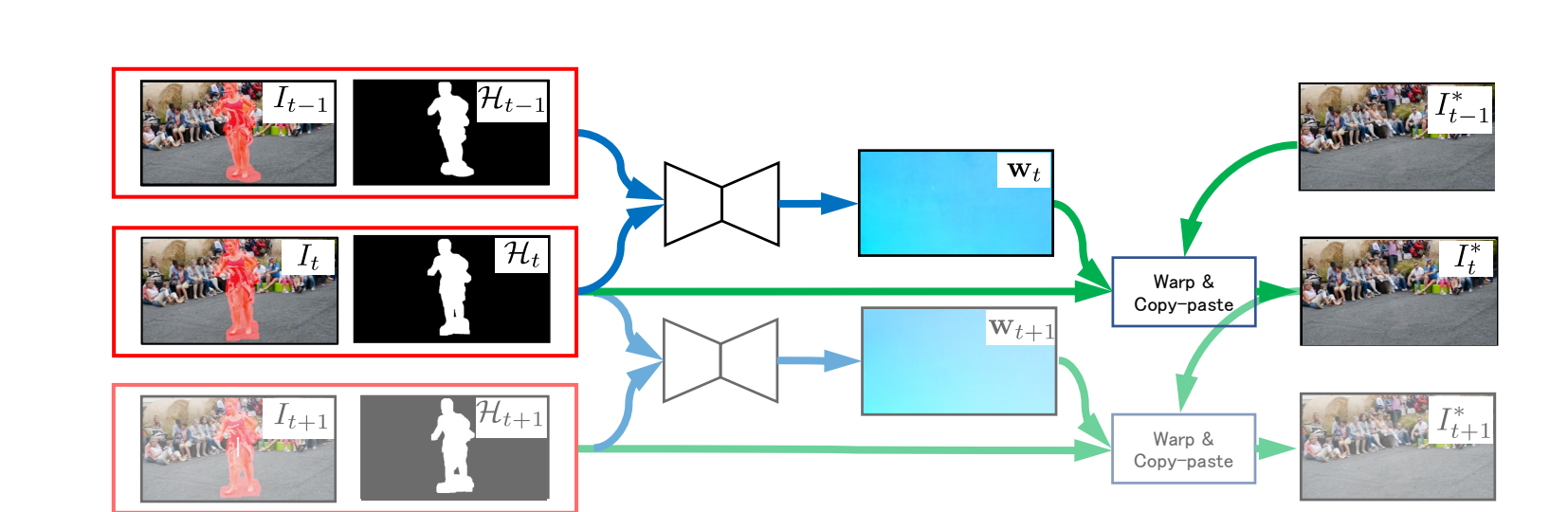
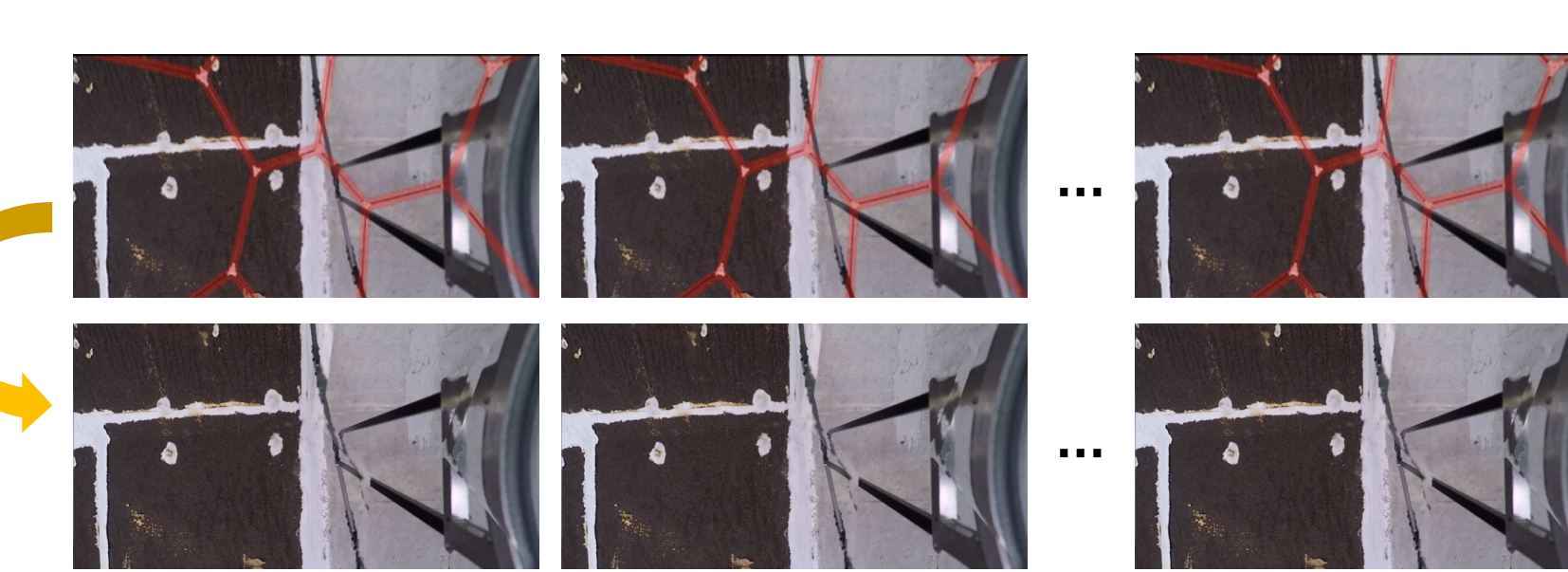
ドローンハードウェア [Salaan et al. ICRA2019他]

- 小型化・面に対する安定多点接触・有線給電等外部アクセスなどの要請に応えるため、ドローン全体ではなく個々のロータを球殻で保護する方式を新たに開発

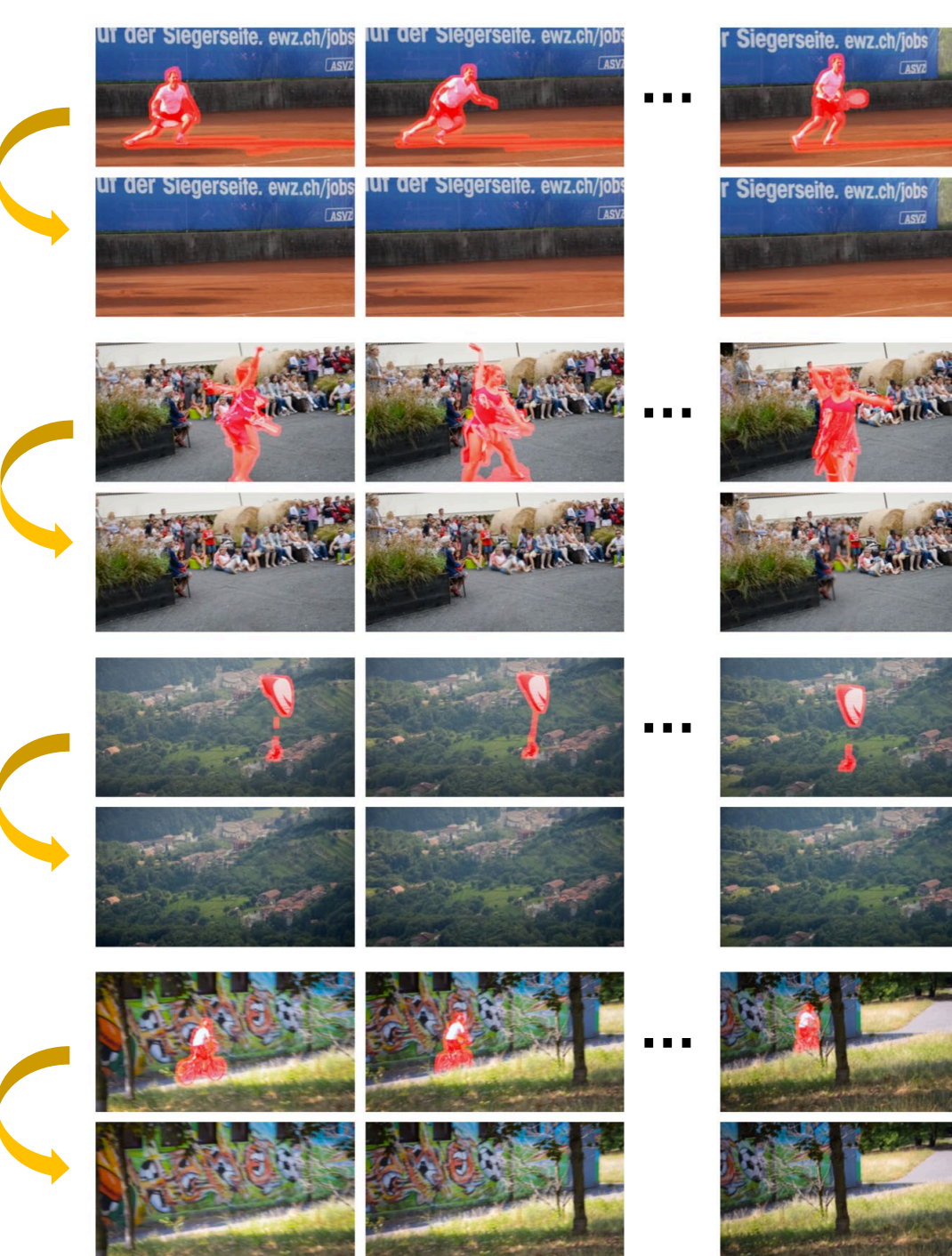


ワイヤフレームのリアルタイム除去 [WACV2019]

- 画像中の球殻ドローンのワイヤフレームを除去
- 世界最高速・30フレーム/秒の速度 ※100万画素未満の場合
- 深層ニューラルネットの利用で実時間処理可能

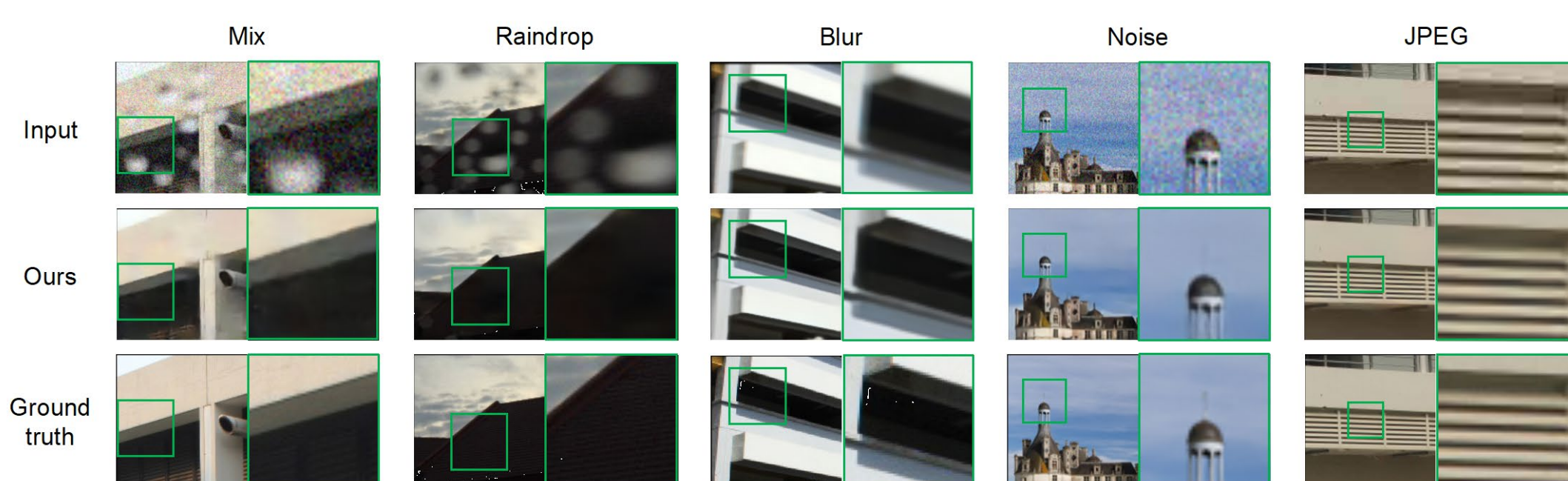


一般的なビデオでの実行例

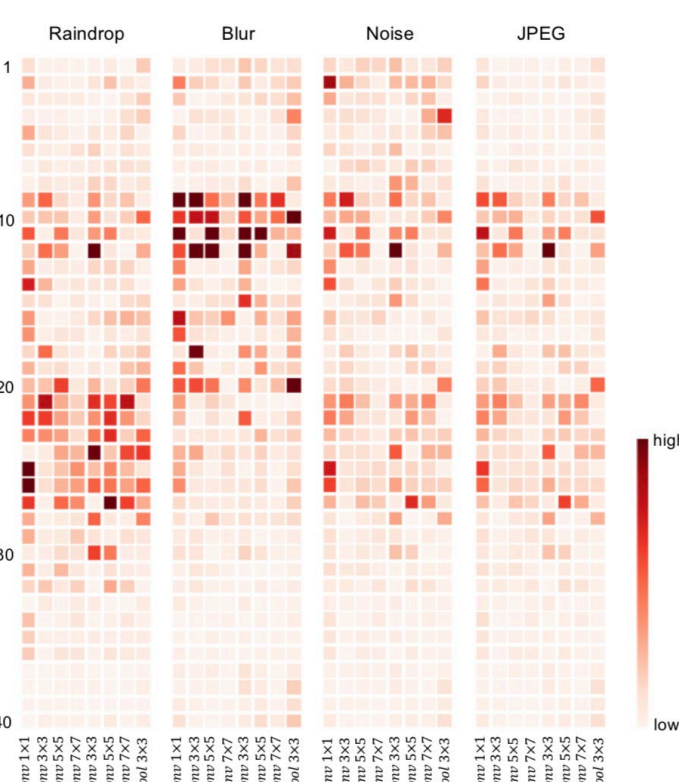


複合的劣化を伴う画像の画質改善 [Suganuma-Liu-Okatani, CVPR2019]

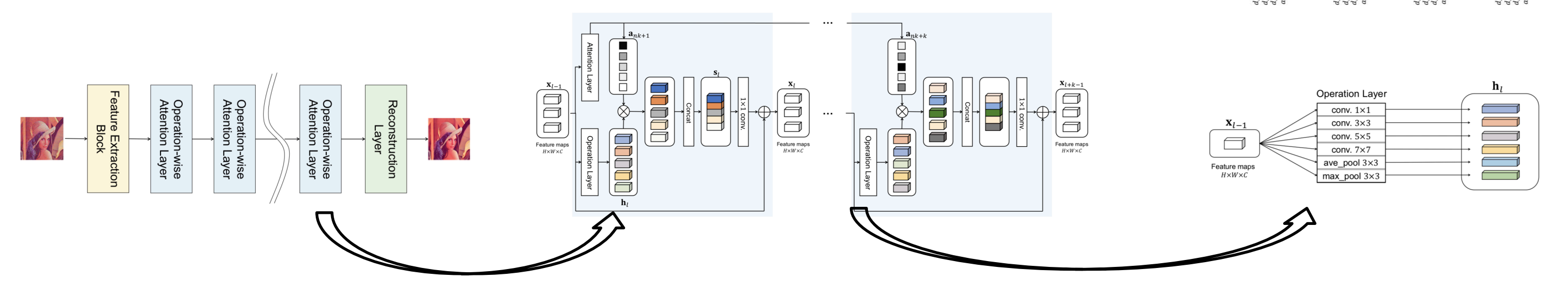
- 事前に画質劣化の要因が不明な場合でも、画質を改善可能な方法を実現



劣化要因に応じた演算の選択の様子



- 入力画像の劣化要因に応じて適応的に演算を選択
- 注意機構を利用したネットワーク構造



画像と言語の表現のマルチタスク学習 [Kien-Okatani, CVPR2019]

- 画像と言語のマルチモーダルタスクのマルチタスク学習に世界で初めて成功
- 複数タスクを1つのネットワークで学習することで、相乗効果を発揮 → 性能向上
- 訓練データをタスク間で共有できるためデータセット生成のコストを下げられる
- タスク間の階層構造に着目、これに対応する層構造をもつネットワーク(5層構造)を設計

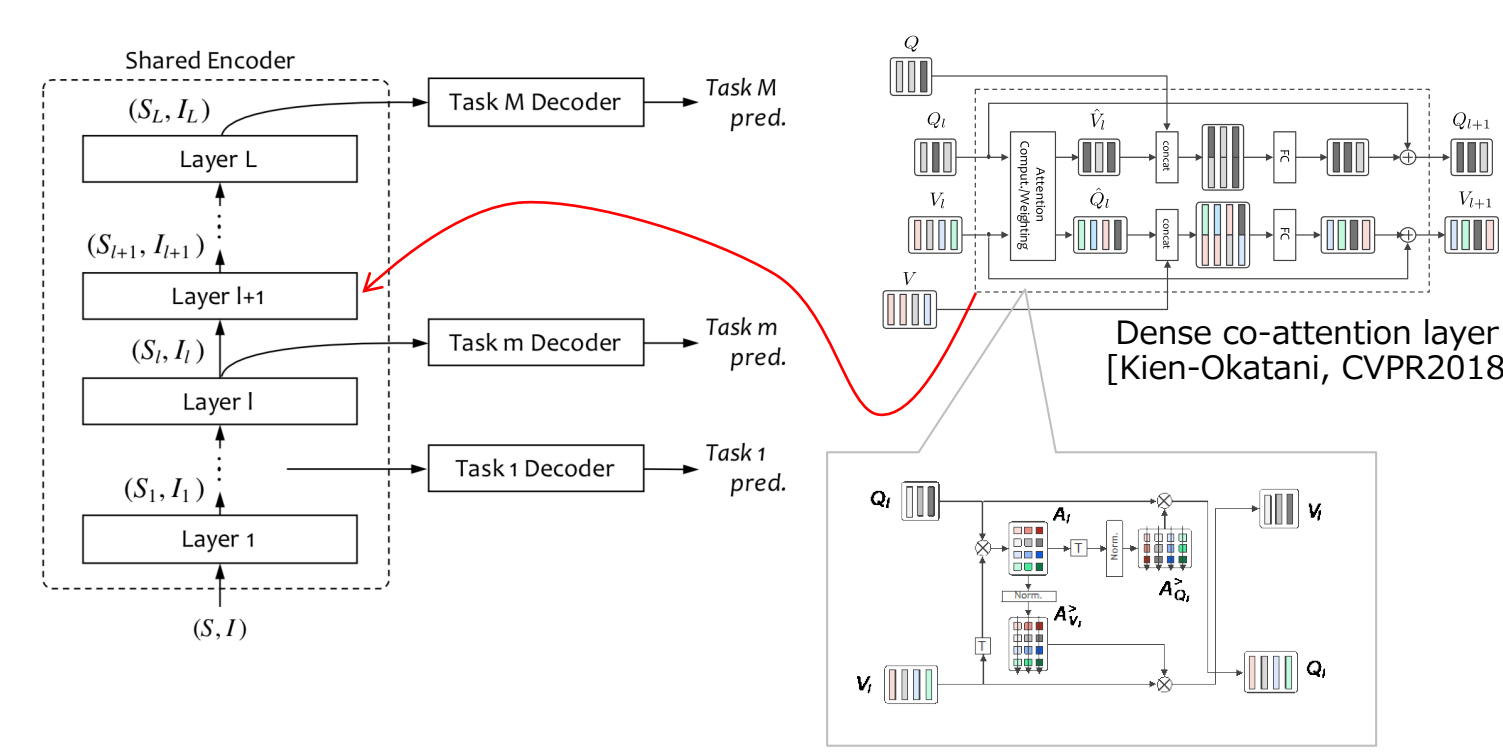


Image Caption Retrieval

Flickr30K, MS-COCO

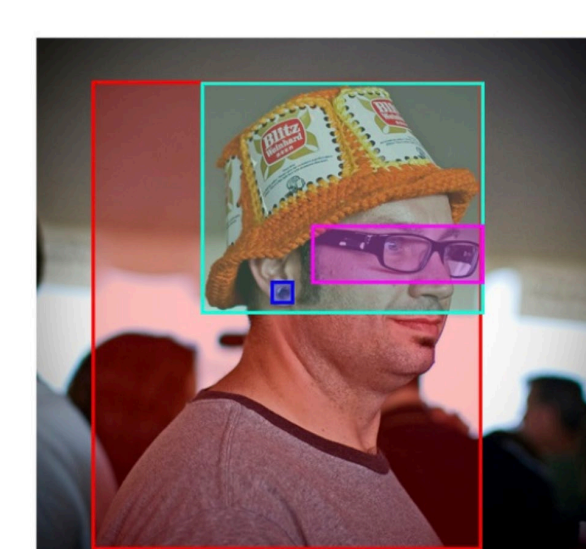
Visual Grounding

Flickr30k Entities

Visual Question Answering

VQA, VQA2.0

What is the mustache made of?



A man with pierced ears is wearing glasses and an orange hat.

A man with glasses is wearing a blue hat and an orange hat.

A man in an orange hat is wearing a blue hat.

A man wears an orange hat and glasses.

Table 3: Results of image annotation and retrieval on the Flickr30K and MSCOCO (1000 testing) datasets.

Method	Flickr30K dataset						MSCOCO dataset					
	R@1	R@5	R@10	R@1	R@5	R@10	R@1	R@5	R@10	R@1	R@5	R@10
RNN+V [19]	34.7	62.7	72.6	26.2	55.1	69.2	40.8	71.9	83.2	29.6	64.8	80.5
OEM [37]	-	-	-	46.7	78.6	88.9	37.9	73.7	85.9	-	-	-
VQA [33]	33.9	62.5	74.5	24.9	52.6	64.8	50.5	80.1	89.7	37.0	70.9	82.9
RTP [31]	37.4	63.1	74.5	26.0	56.0	69.3	-	-	-	-	-	-
DSPE [38]	40.3	68.9	79.9	29.7	60.1	72.1	50.1	79.7	89.2	39.6	75.2	86.9
sm-LSTM [15]	42.5	71.9	81.5	30.2	60.4	72.3	53.2	83.1	91.5	40.7	75.8	87.4
RFP [31]	47.6	77.4	87.1	35.4	68.3	79.9	56.4	85.3	91.5	43.9	78.1	88.6
2WayNet [9]	49.8	67.5	-	36.0	55.6	-	55.8	75.2	-	39.7	63.3	-
DAN [28]	55.0	81.8	89.0	39.4	69.2	79.1	-	-	-	-	-	-
VSE++ [11]	53.9	79.1	87.2	39.6	69.6	79.5	64.6	89.1	95.7	52.0	83.1	92.0
S-E Model [10]	55.5	82.0	89.3	41.1	70.5	80.1	69.9	92.9	97.5	56.7	87.5	94.8
Ours	71.6	84.6	90.8	56.1	82.9	89.4	70.2	89.2	95.9	57.4	88.4	95.6